

1 **IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

2
3 Date: May 2, 2008

4
5 In re application of:

6 **Kelkar et al**

7 Serial No.: 10/629,448

8 Filed: July 29, 2003

9 Group Art Unit: 1631

10 Examiner: **Loria Clow**

11 FOR: **Method and Program**

12 **Product for Discovering**

13 **Similar Gene Expression Profiles**

I hereby certify that this correspondence is being EFS-Web
or facsimile transmitted to the United States Patent and
Trademark Office, Fax No. (571) 273-8300

on: 5 MAY 2008 by: KARL O. HESSE

Karl O. Hesse
Signature

4 MAY 2008
Date of Signature

14
15 **THIRD AMENDED APPEAL BRIEF IN SUPPORT OF APPEAL FROM**
16 **THE PRIMARY EXAMINER TO THE BOARD OF APPEALS**

17
18 Assistant Commissioner for Patents
19 Washington DC 20231

20
21 Sir:

22 Appellants herewith submit a Third Amended Appeal Brief in
23 support of the appeal to the Board of Patent Appeals and
24 Interferences from the decision dated May 16, 2007 of the Primary
25 Examiner finally rejecting claims 1-6, 10-16 and 20.

26
27 **(I) Real Party in Interest**

28 The real party in interest in this appeal is International
29 Business Machines Corporation, a New York corporation, assignee of
30 the entire right, title and interest in the claimed invention.

31
32
Serial No.: 10/629,448

1 **(II) Related Appeals and Interferences**

2 No other appeals or interferences are known to the Appellants,
3 the Appellants' legal representative, or assignee that will directly
4 affect or be directly affected by or have a bearing on the Board's
5 decision in this appeal.

6
7 **(III) Status of Claims**

8 Claims 1-6, 10-16 and 20 are pending in this application.

9 Claims 7-9 and 17-19 were canceled after restriction.

10 The rejection of claims 1-6, 10-16 and 20 under 35 U.S.C. 101
11 is appealed.

12
13
14 **(IV) Status of Amendments**

15
16 The amendment filed after final has not been entered.

1 **(V) Summary of Claimed Subject Matter**

2
3 References to paragraphs are made to the application as
4 published.

5
6 1. A method for determining similarity between portions of gene
7 expression profiles in a computer comprising the steps of:

8 processing a number of gene expression profiles with a similar
9 sequences algorithm that is a time and intensity invariant
10 correlation function to obtain a data set of gene expression profile
11 pairs and a match fraction for each gene expression profile pair;

12 **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

13 listing gene expression profile pairs in clusters by their
14 match fractions; **Page 8 Lines 16-26 (paragraph 38) and Fig.2 # 223**

15 removing a first gene expression profile from a cluster when
16 another cluster has another gene expression profile with a higher
17 match fraction with the first gene expression profile, unless the
18 another gene expression profile requires a larger number of
19 subsequences to achieve similarity with the first gene expression
20 profile; **Page 8 Line 27-Page 9 Line 7 (paragraph 39) and Fig. 2 #**

21 **225**

22 repeating the removing step until all gene expression profiles
23 are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7**

24 **(Paragraph 39) last sentence**

25 providing output of the listing of clusters of gene
26 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**

27 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 2. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:

3 processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function to obtain a data set of gene expression pairs
6 and a match fraction for each pair; **Page 7,Line 31 - Page 8,Line 9**
7 **(Paragraph 36) and Fig. 1, # 213, 215)**

8 listing gene expression pairs in clusters by their match
9 fractions; **Page 8 Lines 16-26 (Paragraph 38) and Fig.2 # 223**

10 removing a first gene from a first cluster when the first gene
11 is also in a second cluster which has another gene with a higher
12 match fraction with the first gene than any of the genes in the
13 first cluster have with the first gene, but; **Page 8 Line 27 - Page 9**
14 **Line 7 (Paragraph 39) and Fig. 2 # 225**

15 retaining the first gene in the first cluster and removing the
16 first gene from the second cluster when the difference between the
17 highest match fraction of the first gene with a gene in the first
18 cluster and the highest match fraction of the first gene with a gene
19 in the second cluster is less than a minimum difference threshold
20 and the number of subsequences represented in the similar gene pair
21 having the highest match fraction in the first cluster is higher
22 than the number of subsequences represented in the similar gene pair
23 having the highest match fraction in the second cluster; **Page 9 Line**
24 **18 - Page 10 Line 10 (Paragraphs 41, 42, 43 and 56) and Fig. 2 # 225**

25 repeating the removing step until all genes are listed in only
26 one cluster; **Page 8 Line 27 - Page 9 Line 7 (Paragraph 39) last**
27 **sentence**

28 providing output of the listing of clusters of gene
29 expression profiles. **Page 6 Lines 24 - 31 and Page 7 Lines 19-24**
30 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 3. A method of determining similarity between portions of gene
2 expression profiles comprising the steps of:

3 processing data embodying a number of gene expression profiles
4 with a similar sequences algorithm in a computer that is a time and
5 intensity invariant correlation function to obtain a data set of
6 gene expression pairs and a match fraction for each pair; **Page**

7 **7,Line 31 - Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

8 choosing a threshold match fraction; **Page 13 Lines 6-20**

9 **(Paragraph 52) sentence 6 and Fig.2 # 217**

10 listing gene expression pairs in clusters by their match
11 fractions above the threshold; **Page 13 Lines 15-20 (Paragraph 52)**
12 **and Fig.2 #s 221, 215, 219 and 223**

13 adding each gene not already in a cluster to a cluster having
14 another gene having a highest match fraction with the each gene
15 without regard of the threshold; **Page 9 Lines 8-14 (Paragraph 40)**
16 **and Fig 2 # 229**

17 removing a first gene from a cluster when the first gene is
18 also in another cluster which has another gene with a higher match
19 fraction with the first gene than any of the genes in the cluster
20 have with the first gene; **Page 8 Line 27-Page 9 Line 7 (Paragraph**
21 **39) first sentence and Fig. 2 # 225**

22 repeating the removing step until all genes are listed in only
23 one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
24 **sentence**

25 providing output of the listing of clusters of gene
26 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
27 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 4. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:

3 processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function with a computer to obtain a data set of gene
6 expression pairs and a match fraction for each pair; **Page 7,Line 31-**
7 **Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

8 choosing a threshold match fraction; **Page 13 Lines 5-20**
9 **(Paragraph 52) sentence 6 and Fig.2 # 217**

10 listing gene expression pairs in clusters by their match
11 fractions above the threshold; **Page 8 Lines 16-26 (paragraph 38) and**
12 **Fig.2 # 223**

13 adding each gene not already in a cluster to a cluster having
14 another gene having a highest match fraction disregarding the
15 threshold with the each gene; **Page 9 Lines 8-14 (Paragraph 40) and**
16 **Fig 2 # 229**

17 removing a first gene from a first cluster when the first gene
18 is also in a second cluster which has another gene with a higher
19 match fraction with the first gene than any of the genes in the
20 first cluster have with the first gene, **Page 8 Line 27-Page 9 Line 7**
21 **(Paragraph 39) first sentence and Fig. 2 # 225**

22 but;

23 retaining the first gene in the first cluster and removing the
24 first gene from the second cluster when the difference between the
25 highest match fraction of the first gene with a gene in the first
26 cluster and the highest match fraction of the first gene with a gene
27 in the second cluster is less than a minimum difference threshold
28 and the number of subsequences represented in the similar gene pair
29 having the highest match fraction in the first cluster is higher
30 than the number of subsequences represented in the similar gene pair

1 having the highest match fraction in the second cluster; **Page 8 Line**
2 **27-Page 9 Line 7 (Paragraph 39) second sentence and Fig. 2 # 225**

3
4 repeating the removing and retaining steps until all genes are
5 listed in only one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph**
6 **39) last sentence**

7 providing output of the listing of clusters of gene
8 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
9 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 5. A method in a computer for determining similarity between
2 genes comprising the steps of:

3 listing genes to be compared in a data set by their gene
4 expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36)**
5 **second sentence and Fig. 1 # 211**

6 processing the listed gene expression profiles with a similar
7 sequences algorithm that is a time and intensity invariant
8 correlation function to obtain a data set of gene expression pairs
9 and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9**
10 **(Paragraph 36) and Fig. 1, # 213, 215**

11 choosing a threshold match fraction; **Page 13 Lines 5-20**
12 **(Paragraph 52) sentence 6 and Fig.2 # 217**

13
14 creating a set G in which to list indices of genes accounted
15 for; **Page 8 Lines 11-14 (Paragraph 37) and Fig. 2 # 217**

16 assigning genes i and j to a cluster a if they have a match
17 fraction greater than the threshold; **Page 8 Lines 16-26 (Paragraph**
18 **38) second sentence and Fig. 2 # 223 -> ca={i,j,...}**

19 assigning gene k to the cluster a if it has a match fraction
20 greater than the threshold with either gene i or gene j;
21 **Page 8 Lines 16-26 (Paragraph 38) last sentence and Fig. 2 # 223 ->**
22 **ca={i,j,k,...}**

23 assigning genes k and l to a cluster b if they have a match
24 fraction greater than the threshold and if both gene k and gene l do
25 not have match fractions above the threshold with either gene i or
26 gene j;

27 **Page 8 Lines 16-26 (Paragraph 38) and Fig. 2 # 223 -> cb={k,l,...}**

28 repeating the assigning steps until all genes to be compared
29 have been considered; **Page 13 Lines 5-20 (Paragraph 52) first two**
30 **sentences and Fig #s 217, 219, and 229**

1 removing a first gene from a cluster when another cluster has
2 another gene with a higher match fraction with the first gene; **Page**
3 **8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence and Fig. 2 #**
4 **225**

5 repeating the removing step until all genes are listed in only
6 one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
7 **sentence**

8 providing output of the listing of clusters of gene
9 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
10 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 6. A method in a computer for determining similarity between
2 genes comprising the steps of:

3 listing genes to be compared in a data set by their gene
4 expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36)**
5 **second sentence and Fig. 1 # 211**

6 processing the listed gene expression profiles with a similar
7 sequences algorithm that is a time and intensity invariant
8 correlation function to obtain a data set of gene expression pairs
9 and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9**
10 **(Paragraph 36) and Fig. 1, # 213, 215**

11 choosing a threshold match fraction; **Page 13 Lines 5-20**
12 **(Paragraph 52) sentence 6 and Fig.2 # 217**

13 creating a set G in which to list indices of genes accounted
14 for; **Page 13 Lines 5-20 (Paragraph 52) 3rd and 4th sentences and**
15 **Fig. 2 # 217**

16 assigning genes i and j to cluster 1 if they have a match
17 fraction greater than the threshold; **Page 13 Line 21-Page 14 Line 6**
18 **and Page 14 Lines 8-12 (Paragraphs 53 and 54) first sentence and**
19 **Table IV**

20 assigning gene k to cluster 1 if it has a match fraction
21 greater than the threshold with either gene i or gene j; **Page 13**
22 **Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15**
23 **Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**

24 assigning genes k and l to cluster 2 if they have a match
25 fraction greater than the threshold and if both gene k and gene l do
26 not have match fractions above the threshold with either gene i or
27 gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14**
28 **Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and**
29 **Table IV and Fig. 2 # 225**

30 removing a first gene from a cluster when another cluster has
31 another gene with a higher match fraction with the first gene,

1 unless the another gene requires a larger number of subsequences to
2 achieve similarity with the first gene; **Page 14 Line 24 - Page 15**
3 **Line 3 (Paragraph 56) beginning at the 3rd sentence and Fig. 2 # 225**
4 repeating the removing step until all genes are listed in only
5 one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
6 **sentence**
7 providing output of the listing of clusters of gene
8 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
9 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 10. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;

8 **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

9 programmed means for listing gene expression pairs in clusters
10 by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and**
11 **Fig.2 # 223**

12 programmed means for removing a first gene from a cluster when
13 the first gene is also in another cluster which has another gene
14 with a higher match fraction with the first gene than any of the
15 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
16 **Line 7 (paragraph 39) and Fig. 2 # 225**

17 programmed means for repeating the removing step until all
18 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
19 **(Paragraph 39) last sentence**

1 11. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles using output from a similar sequences algorithm
4 that is a time and intensity invariant correlation function
5 comprising:

6 programmed means for providing a gene expression profile data
7 set as input to programmed means embodying a similar sequences
8 algorithm that is a time and intensity invariant correlation
9 function to obtain a data set of gene expression pairs and a match
10 fraction for each pair as output from the programmed means embodying
11 a similar sequences algorithm; **Page 7,Line 31-Page 8,Line 9; Page**
12 **10, Line 22-Page 11,Line 7 (Paragraphs 36, 44) and Fig. 1, # 213,**
13 **215)**

14 programmed means for listing the gene expression pairs in
15 clusters by their match fractions; **Page 7,Line 31-Page 8,Line 9**
16 **(Paragraph 36) beginning at the second sentence and Fig. 1 # 215**

17
18 programmed means for removing a first gene from a cluster when
19 the first gene is also in another cluster which has another gene
20 with a higher match fraction with the first gene than any of the
21 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
22 **Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**

23 programmed means for repeating the removing step until all
24 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
25 **(Paragraph 39) last sentence**

1 12. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair; **Page**
8 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

9 programmed means for listing gene expression pairs in clusters
10 by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and**
11 **Fig.2 # 223**

12 programmed means for removing a first gene from a first cluster
13 when the first gene is also in a second cluster which has another
14 gene with a higher match fraction with the first gene than any of
15 the genes in the first cluster have with the first gene, **Page 8 Line**
16 **27-Page 9 Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**
17 but;

18 programmed means for retaining the first gene in the first
19 cluster and removing the first gene from the second cluster when the
20 difference between the highest match fraction of the first gene with
21 a gene in the first cluster and the highest match fraction of the
22 first gene with a gene in the second cluster is less than a minimum
23 difference threshold and the number of subsequences represented in
24 the similar gene pair having the highest match fraction in the first
25 cluster is higher than the number of subsequences represented in the
26 similar gene pair having the highest match fraction in the second
27 cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence**
28 **and Fig. 2 # 225**

29 programmed means for repeating the removing step until all
30 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
31 **(Paragraph 39) last sentence**

1 13. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair; **Page**
8 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

9 programmed means for choosing a threshold match fraction; **Page**
10 **13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

11 programmed means for listing gene expression pairs in clusters
12 by their match fractions above the threshold; **Page 8 Lines 16-26**
13 **(paragraph 38) and Fig.2 # 223**

14 programmed means for adding each gene not already in a cluster
15 to a cluster having another gene having a highest match fraction
16 with the each gene without regard of the threshold; **Page 9 Lines 8-**
17 **14 (Paragraph 40) and Fig 2 # 229**

18 programmed means for removing a first gene from a cluster when
19 the first gene is also in another cluster which has another gene
20 with a higher match fraction with the first gene than any of the
21 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
22 **Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**

23 programmed means for repeating the removing step until all
24 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
25 **(Paragraph 39) last sentence**

1 14. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair; **Page**
8 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

9 programmed means for choosing a threshold match fraction; **Page**
10 **13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

11 programmed means for listing gene expression pairs in clusters
12 by their match fractions above the threshold; **Page 8 Lines 16-26**
13 **(paragraph 38) and Fig.2 # 223**

14 programmed means for adding each gene not already in a cluster
15 to a cluster having another gene having a highest match fraction
16 disregarding the threshold with the each gene; **Page 9 Lines 8-14**
17 **(Paragraph 40) and Fig 2 # 229**

18 programmed means for removing a first gene from a first cluster
19 when the first gene is also in a second cluster which has another
20 gene with a higher match fraction with the first gene than any of
21 the genes in the first cluster have with the first gene, **Page 8 Line**
22 **27-Page 9 Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**
23 but;

24 programmed means for retaining the first gene in the first
25 cluster and removing the first gene from the second cluster when the
26 difference between the highest match fraction of the first gene with
27 a gene in the first cluster and the highest match fraction of the
28 first gene with a gene in the second cluster is less than a minimum
29 difference threshold and the number of subsequences represented in
30 the similar gene pair having the highest match fraction in the first
31 cluster is higher than the number of subsequences represented in the

1 similar gene pair having the highest match fraction in the second
2 cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence**
3 **and Fig. 2 # 225**
4 programmed means for repeating the removing and retaining steps
5 until all genes are listed in only one cluster. **Page 8 Line 27-Page**
6 **9 Line 7 (Paragraph 39) last sentence**

1 15. A program product having computer readable code stored on a
2 recordable media for determining similarity between genes comprising
3 the steps of:

4 programmed means for listing genes to be compared by their gene
5 expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36)**
6 **second sentence and Fig. 1 # 211**

7 programmed means for processing the listed gene expression
8 profiles with a similar sequences algorithm that is a time and
9 intensity invariant correlation function to obtain a data set of
10 gene expression pairs and a match fraction for each pair; **Page**
11 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

12 programmed means for choosing a threshold match fraction; **Page**
13 **13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

14 programmed means for creating a null set $G(0)$ to hold genes
15 accounted for; **Page 13 Lines 6-20 (Paragraph 52) 3rd and 4th**
16 **sentences and Fig. 2 # 217**

17 programmed means for assigning genes i and j to cluster 1 if
18 they have a match fraction greater than the threshold; **Page 13 Line**
19 **21-Page 14 Line 6 and Page 14 Lines 8-12 (Paragraphs 53 and 54)**
20 **first sentence and Table IV**

21 programmed means for assigning gene k to cluster 1 if it has a
22 match fraction greater than the threshold with either gene i or gene
23 j ; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line**
24 **24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**

25 programmed means for assigning genes k and l to cluster 2 if
26 they have a match fraction greater than the threshold and if both
27 gene k and gene l do not have match fractions above the threshold
28 with either gene i or gene j ; **Page 13 Line 21-Page 14 Line 6, Page**
29 **14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd**
30 **sentence, 56) and Table IV and Fig. 2 # 225**

1 programmed means for removing a first gene from a cluster when
2 another cluster has another gene with a higher match fraction with
3 the first gene; **Page 14 Line 24 - Page 15 Line 3 (Paragraph 56)**

4 **beginning at the 3rd sentence and Fig. 2 # 225**

5 programmed means for repeating the removing step until all
6 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
7 **(Paragraph 39) last sentence**

1 16. A program product having computer readable code stored on a
2 recordable media for determining similarity between genes comprising
3 the steps of:

4 programmed means for listing genes to be compared by their gene
5 expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36)**
6 **second sentence and Fig. 1 # 211**

7 programmed means for processing the listed gene expression
8 profiles with a similar sequences algorithm that is a time and
9 intensity invariant correlation function to obtain a data set of
10 gene expression pairs and a match fraction for each pair; **Page**
11 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

12 programmed means for choosing a threshold match fraction; **Page**
13 **13 Lines 5-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

14 programmed means for creating a null set $G(0)$ to hold genes
15 accounted for; **Page 8 Lines 11-14 (Paragraph 37 and Fig. 2 # 217**

16 programmed means for assigning genes i and j to cluster 1 if
17 they have a match fraction greater than the threshold; **Page 13 Line**
18 **21-Page 14 Line 6 and Page 14 Lines 8-12 (Paragraphs 53 and 54)**
19 **first sentence and Table IV**

20 programmed means for assigning gene k to cluster 1 if it has a
21 match fraction greater than the threshold with either gene i or gene
22 j ; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line**
23 **24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**

24 programmed means for assigning genes k and l to cluster 2 if
25 they have a match fraction greater than the threshold and if both
26 gene k and gene l do not have match fractions above the threshold
27 with either gene i or gene j ; **Page 13 Line 21-Page 14 Line 6, Page**
28 **14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd**
29 **sentence, 56) and Table IV and Fig. 2 # 225**

1 programmed means for removing a first gene from a cluster when
2 another cluster has another gene with a higher match fraction with
3 the first gene, unless the another gene requires a larger number of
4 subsequences to achieve similarity with the first gene; **Page 14 Line**
5 **24 - Page 15 Line 3 (Paragraph 56) beginning at the 3rd sentence and**
6 **Fig. 2 # 225**

7 programmed means for repeating the removing step until all
8 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
9 **(Paragraph 39) last sentence**

1 20. In a method of determining similarity between portions of
2 gene expression profiles which includes processing a number of gene
3 expression profiles using a computer with a similar sequences
4 algorithm that is a time and intensity invariant correlation
5 function to obtain a data set of gene expression pairs and a match
6 fraction for each pair, **Page 7,Line 31-Page 8,Line 9 (Paragraph 36**
7 **and Fig. 1, # 213, 215** the improvement comprising the steps of:

8 listing gene expression pairs in clusters by their match
9 fractions; **Page 8 Lines 16-26 (paragraph 38) and Fig.2 # 223**

10 removing a first gene from a cluster when another cluster has
11 another gene with a higher match fraction with the first gene,
12 unless the another gene requires a larger number of subsequences to
13 achieve similarity with the first gene; **Page 14 Line 24 - Page 15**
14 **Line 3 (Paragraph 56) beginning at the 3rd sentence and Fig. 2 # 225**

15 repeating the removing step until all genes are listed in only
16 one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
17 **sentence**

18 providing output of the listing of clusters of gene
19 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
20 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1 **(VI) Grounds of Rejection to be reviewed on Appeal**

2
3 Claims 1-6, 10-16 and 20 are pending in this application.

4
5 Claims 7-9 and 17-19 were canceled after restriction.

6
7 The rejection of claims 1-6, 10-16 and 20 under 35 U.S.C. 101
8 is appealed.

9
10 The issues in this appeal are:

11
12 I. Whether claims 1, 2, 4, 6, 12, 14, 16 and 20 are indefinite
13 under 35 U.S.C.112 second paragraph on the ground that the
14 specification does not provide a meaning for the word
15 "subsequences". The Examiner has found Figures 4 and 6 and
16 paragraphs 48 - 50 to be unclear. It is believed that this issue
17 has been resolved by applicants' remarks in the amendment after
18 final.

19
20 II. Whether output to a user is a required claim step in order to
21 define an invention, that is a practical application which is
22 useful, concrete and tangible.

23 Claim 10 is representative of claims 1-6, 10-16 and 20 which
24 have been rejected under 35 U.S.C. 101 and is related to Issue II.

1 **(VII) Argument**

2
3 The rejection of claims 1, 2, 4, 6, 12, 14, 16 and 20 under 35
4 U.S.C. 112 has been addressed by applicant's attorney in the
5 amendment after final and the Examiners remarks indicate that the
6 Examiner has found applicants' arguments to be persuasive, therefore
7 it was believed that this issue no longer existed, however in view
8 of the Examiners comment in the Notice of Non-Compliant Appeal
9 Brief, these remarks are repeated below as Issue I.
10

11 When this application was filed in 2003, applicants' attorney
12 believed that providing independent claims would facilitate
13 prosecution because they could be allowed or rejected without
14 requiring rewriting to incorporate independent claim limitations
15 into dependent claims when they were found to be allowable. It is
16 understood that currently this practice is not permitted in excess
17 of four claims and applicants' attorney regrets any inconvenience
18 caused thereby.
19
20
21

22 **Issue I** Whether claims 1, 2, 4, 6, 12, 14, 16 and 20 are
23 indefinite under 35 U.S.C.112 second paragraph on the ground that
24 the specification does not provide a meaning for the word
25 "subsequences". The Examiner has found Figures 4 and 6 and
26 paragraphs 48 - 50 to be unclear.
27

28 The specification at paragraph 15 recites:

29 [0015] According to the instant invention, two temporal sequences are similar and can be placed in
30 the same cluster if they have enough non-overlapping time-ordered pairs of sub-sequences that are
31 similar.

32 Paragraph 50 recites:
Serial No.: 10/629,448

1 [0050] FIG. 4 exemplifies noise resistance and partial similarity. When one looks at *gene* 4 and *gene*
2 3, it is clear that most likely, the value of 10 for *gene* 3 at t=3 is an outlier. This data point could have
3 occurred, either from manual error or instrumentation error. The Agrawal Fast Similarity Search
4 algorithm will minimize this artifact data point by its design, and identify two matching areas. The
5 profile from t=1 to t=2 is identified as one subsequence and the profile from t=4 to t=6 as another
6 subsequence. Since it has minimized this "outlier or noise", it is able to identify these two genes as
7 similar in function.

8
9 Applicants believe that it is clear to those skilled in the art that
10 the genes 3 and 4 in Figure 4 are temporal sequences from time 1
11 through time 6 and that each sequence includes subsequences from
12 time 1 to time 2 and from time 4 to time 6.

13
14 In view of the Examiners finding that the above remarks presented in
15 the amendment after final were persuasive, it is believed that this
16 issue has been resolved and would be removed if the amendment after
17 final were entered.

1
2 **Issue II:** Whether output to a user is a required claim step in order
3 to define an invention, that is a practical application which is
4 useful, concrete and tangible.

5
6 Appellants claim in representative claim 10:

7 10. A program product having computer readable code stored on a
8 recordable media for determining similarity between portions of gene
9 expression profiles comprising:

10 programmed means for processing a number of gene expression
11 profiles with a similar sequences algorithm that is a time and
12 intensity invariant correlation function to obtain a data set of
13 gene expression pairs and a match fraction for each pair; **Page**
14 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

15 programmed means for listing gene expression pairs in clusters
16 by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and**
17 **Fig.2 # 223**

18 programmed means for removing a first gene from a cluster when
19 the first gene is also in another cluster which has another gene
20 with a higher match fraction with the first gene than any of the
21 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
22 **Line 7 (paragraph 39) and Fig. 2 # 225**

23 programmed means for repeating the removing step until all
24 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
25 **(Paragraph 39) last sentence**

1 Applicants believe that the specification and claims indeed do
2 describe a method and a program product that produce a result that
3 has substantial and credible utility as required by MPEP 2107 II and
4 that the claims are limited to a narrow practical application in a
5 computer related art.

6
7 The Examiner relies on the "New Interim Guidelines" to
8 interpret the requirements of the Federal Courts under the current
9 law to require claiming "output to a user". Applicants believe that
10 the Examiner is mistaken and is applying an interpretation of the
11 definition of the word tangible that is:

- 12 1) narrower than appropriate under the current law and is
13 2) narrower than required under the "New Guidelines".
14

15 1) The introduction to the "New Guidelines" states:
16 "These Examination Guidelines ("Guidelines") are based on the USPTO's current understanding of
17 the law and are believed to be fully consistent with binding precedent of the Supreme Court, the
18 Federal Circuit and the Federal Circuit's predecessor courts. These Guidelines do not constitute
19 substantive rulemaking and hence do not have the force and effect of law."
20

21 In following the "Guidelines", the Examiner appears to require
22 separate interpretations of the words useful, concrete and tangible.
23

24 Applicants' attorney has found no basis in any of the Federal
25 Circuit opinions using these words that imply that these terms are
26 to have separate meanings. They appear to always be used together
27 as synonyms for the concept of being useful and non-abstract.
28 Applicants' attorney has requested that the Examiner provide a
29 citation to a court's requirement that these terms are part of a
30 three pronged test if such is the case in order to help applicants
31 decide whether to appeal or request continued examination. No
32 citation was provided.

1 2) Even under the "Guidelines, the Examiners interpretation of
2 the word tangible is unnecessarily narrow.
3 The "Guidelines" at page 13 recite
4 "Accordingly, a complete definition of the scope of 35 U.S.C. § 101, reflecting Congressional intent,
5 is that any new and useful process, machine, manufacture or composition of matter under the sun that
6 is made by man is the proper subject matter of a patent. The subject matter courts have found to be
7 outside of, or exceptions to, the four statutory categories of invention is limited to abstract ideas, laws
8 of nature and natural phenomena. While this is easily stated, determining whether an applicant is
9 seeking to patent an abstract idea, a law of nature or a natural phenomenon has proven to be
10 challenging."

11
12 Beginning at page 21 the "Guidelines" recite:

13 **"TANGIBLE RESULT"**

14 "The tangible requirement does not necessarily mean that a claim must either be tied to a particular
15 machine or apparatus or must operate to change articles or materials to a different state or thing. However,
16 the tangible requirement does require that the claim must recite more than a § 101 judicial exception, in
17 that the process claim must set forth a practical application of that § 101 judicial exception to produce a
18 real-world result. Benson, 409 U.S. at 71-72, 175 USPQ at 676-77 (invention ineligible because had "no
19 substantial practical application."). "[A]n application of a law of nature or mathematical formula to a ...
20 process may well be deserving of patent protection." Diehr, 450 U.S. at 187, 209 USPQ at 8 (emphasis
21 added); see also Corning, 56 U.S. (15 How.) at 268, 14 L.Ed. 683 ("It is for the discovery or invention of
22 some practical method or means of producing a beneficial result or effect, that a patent is granted . . .").

23
24 In other words, the opposite meaning of "tangible" is "abstract."
25 The bare conversion of any binary data as in Gottschalk V. Benson or
26 the bubble sort of any data as in "Warmerdam, 33 F.3d at 1360, 31 USPQ2d at 1759
27 ("steps of 'locating' a medial axis, and 'creating' a bubble hierarchy ... describe nothing more than the
28 manipulation of basic mathematical constructs, the paradigmatic 'abstract idea'")" recited at
29 page 14 of the "Guidelines" are examples of the abstract.

1 Applicants' process does not convert or process just any data but is
2 limited to useful concrete and non-abstract gene expression profiles
3 in a data base of such profiles. Applicants' process is but one
4 application of many possible applications of the mathematical steps
5 involved in obtaining the useful result.
6

7 At page 17 of the "Guidelines we see:

8 While abstract ideas, natural phenomena, and laws of nature are not eligible for patenting, methods and
9 products employing abstract ideas, natural phenomena, and laws of nature to perform a real-world
10 function may well be. In evaluating whether a claim meets the requirements of section 101, the claim
11 must be considered as a whole to determine whether it is for a particular application of an abstract idea,
12 natural phenomenon, or law of nature, rather than for the abstract idea, natural phenomenon, or law of
13 nature itself.
14

15 As is clear from the specification and the claim limitations,
16 applicants' process is limited to a particular practical application
17 and is not an abstract idea, natural phenomenon or a law of nature.
18

19 The result is that all of the processed gene expression profiles are
20 each listed in only one cluster. This result of applicants' claims
21 is a very useful, repeatable and non-abstract result which is
22 recognized by those skilled in the medical and computer arts to be
23 of great value and useful, non-abstract and concrete finding of
24 similar gene expression profiles.
25

1 **PRIOR ART**

2 Applicants note that their claims have not been rejected on prior
3 art yet have been restricted on the ground that there were two
4 groups of claims that required two fields of search. It is not
5 apparent whether relevant prior art patents were considered by the
6 Examiner while examining this application. It is believed that the
7 "Guidelines" on page 10 are helpful in determining both the novelty
8 of applicants' invention and the **usefulness and non-abstract nature**
9 of applicants' the invention.

10
11 As evidenced by the references which applicants have attempted to
12 incorporate by reference, but have acquiesced to the Examiners
13 correct requirement to cancel, in addition to applicants teachings
14 in the background art section of their specification, users in the
15 medical profession find great value and usefulness in methods for
16 finding similar gene expression profiles that are tangible and
17 concrete. See for example US Patent 6,406,853 abstract and claims
18 25, 26 and US Patent 6,436,642 column 26 beginning at line 15.

19
20 It is believed that if the rejections under 35 U.S.C. 101 put forth
21 in this application were appropriate, many of the relevant prior art
22 patents in the appropriate fields of search would be found to be
23 invalid. Since they were issued under the guidance of current
24 statutory law and court cases, it must be that the rejections in
25 this application are based upon excessively narrow and untenable
26 interpretation of the current law.

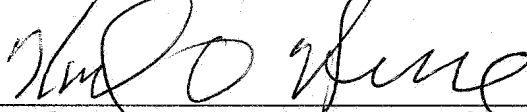
1
2
3 Accordingly it is believed that the claims are clear, statutory
4 and definite and are drawn to a novel and unobvious method and
5 program product for clustering gene expression profiles which result
6 is concrete, tangible and directly useful in drug selection and
7 disease diagnosis.

8
9 When applicants' amendment after final or an equivalent
10 amendment is entered, the outstanding objection to the specification
11 will be removed as applicants have therein canceled the attempted
12 incorporation of US Patents 6,406,853 and 6,436,642 by reference.

13
14
15 **Request for Relief**

16
17 Wherefore, Appellants respectfully request that the rejection
18 of pending claims 1 - 6, 10 - 16 and 20 be reversed.
19
20

21 Respectfully submitted,

22 
23
24

25 Karl O. Hesse, Reg. No. 25,398

26 Attorney for Appellants

27 IBM Corporation
28 Intellectual Property Law
29 MG90-201/1
30 8501 IBM Drive
31 Charlotte, NC 28262-8563

Land line (704) 895-8241

Cell phone (704) 724-1413

Fax: (704) 594-8307

1 **(VIII) Appendix**
2 **Claims Involved in this Appeal**
3

4 1. A method for determining similarity between portions of gene
5 expression profiles in a computer comprising the steps of:

6 processing a number of gene expression profiles with a similar
7 sequences algorithm that is a time and intensity invariant
8 correlation function to obtain a data set of gene expression profile
9 pairs and a match fraction for each gene expression profile pair;

10 listing gene expression profile pairs in clusters by their
11 match fractions;

12 removing a first gene expression profile from a cluster when
13 another cluster has another gene expression profile with a higher
14 match fraction with the first gene expression profile, unless the
15 another gene expression profile requires a larger number of
16 subsequences to achieve similarity with the first gene expression
17 profile;

18 repeating the removing step until all gene expression profiles
19 are listed in only one cluster;

20 providing output of the listing of clusters of gene
21 expression profiles.
22

1 2. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:
3 processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function to obtain a data set of gene expression pairs
6 and a match fraction for each pair;
7 listing gene expression pairs in clusters by their match
8 fractions;
9 removing a first gene from a first cluster when the first gene
10 is also in a second cluster which has another gene with a higher
11 match fraction with the first gene than any of the genes in the
12 first cluster have with the first gene, but;
13 retaining the first gene in the first cluster and removing the
14 first gene from the second cluster when the difference between the
15 highest match fraction of the first gene with a gene in the first
16 cluster and the highest match fraction of the first gene with a gene
17 in the second cluster is less than a minimum difference threshold
18 and the number of subsequences represented in the similar gene pair
19 having the highest match fraction in the first cluster is higher
20 than the number of subsequences represented in the similar gene pair
21 having the highest match fraction in the second cluster;
22 repeating the removing step until all genes are listed in only
23 one cluster;
24 providing output of the listing of clusters of gene
25 expression profiles.

1 3. A method of determining similarity between portions of gene
2 expression profiles comprising the steps of:
3 processing data embodying a number of gene expression profiles
4 with a similar sequences algorithm in a computer that is a time and
5 intensity invariant correlation function to obtain a data set of
6 gene expression pairs and a match fraction for each pair;
7 choosing a threshold match fraction;
8 listing gene expression pairs in clusters by their match
9 fractions above the threshold;
10 adding each gene not already in a cluster to a cluster having
11 another gene having a highest match fraction with the each gene
12 without regard of the threshold;
13 removing a first gene from a cluster when the first gene is
14 also in another cluster which has another gene with a higher match
15 fraction with the first gene than any of the genes in the cluster
16 have with the first gene;
17 repeating the removing step until all genes are listed in only
18 one cluster;
19 providing output of the listing of clusters of gene
20 expression profiles.

1 4. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:
3 processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function with a computer to obtain a data set of gene
6 expression pairs and a match fraction for each pair;
7 choosing a threshold match fraction;
8 listing gene expression pairs in clusters by their match
9 fractions above the threshold;
10 adding each gene not already in a cluster to a cluster having
11 another gene having a highest match fraction disregarding the
12 threshold with the each gene;
13 removing a first gene from a first cluster when the first gene
14 is also in a second cluster which has another gene with a higher
15 match fraction with the first gene than any of the genes in the
16 first cluster have with the first gene, but;
17 retaining the first gene in the first cluster and removing the
18 first gene from the second cluster when the difference between the
19 highest match fraction of the first gene with a gene in the first
20 cluster and the highest match fraction of the first gene with a gene
21 in the second cluster is less than a minimum difference threshold
22 and the number of subsequences represented in the similar gene pair
23 having the highest match fraction in the first cluster is higher
24 than the number of subsequences represented in the similar gene pair
25 having the highest match fraction in the second cluster;
26 repeating the removing and retaining steps until all genes are
27 listed in only one cluster;
28 providing output of the listing of clusters of gene
29 expression profiles.

1 5. A method in a computer for determining similarity between
2 genes comprising the steps of:
3 listing genes to be compared in a data set by their gene
4 expression profiles;
5 processing the listed gene expression profiles with a similar
6 sequences algorithm that is a time and intensity invariant
7 correlation function to obtain a data set of gene expression pairs
8 and a match fraction for each pair;
9 choosing a threshold match fraction;
10 creating a set G in which to list indices of genes accounted
11 for;
12 assigning genes i and j to a cluster a if they have a match
13 fraction greater than the threshold;
14 assigning gene k to the cluster a if it has a match fraction
15 greater than the threshold with either gene i or gene j;
16 assigning genes k and l to a cluster b if they have a match
17 fraction greater than the threshold and if both gene k and gene l do
18 not have match fractions above the threshold with either gene i or
19 gene j;
20 repeating the assigning steps until all genes to be compared
21 have been considered;
22 removing a first gene from a cluster when another cluster has
23 another gene with a higher match fraction with the first gene;
24 repeating the removing step until all genes are listed in only
25 one cluster;
26 providing output of the listing of clusters of gene
27 expression profiles.
28

1 6. A method in a computer for determining similarity between
2 genes comprising the steps of:
3 listing genes to be compared in a data set by their gene
4 expression profiles;
5 processing the listed gene expression profiles with a similar
6 sequences algorithm that is a time and intensity invariant
7 correlation function to obtain a data set of gene expression pairs
8 and a match fraction for each pair;
9 choosing a threshold match fraction;
10 creating a set G in which to list indices of genes accounted
11 for;
12 assigning genes i and j to cluster 1 if they have a match
13 fraction greater than the threshold;
14 assigning gene k to cluster 1 if it has a match fraction
15 greater than the threshold with either gene i or gene j;
16 assigning genes k and l to cluster 2 if they have a match
17 fraction greater than the threshold and if both gene k and gene l do
18 not have match fractions above the threshold with either gene i or
19 gene j;
20 removing a first gene from a cluster when another cluster has
21 another gene with a higher match fraction with the first gene,
22 unless the another gene requires a larger number of subsequences to
23 achieve similarity with the first gene;
24 repeating the removing step until all genes are listed in only
25 one cluster;
26 providing output of the listing of clusters of gene
27 expression profiles.
28

1 10. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;

8 programmed means for listing gene expression pairs in clusters
9 by their match fractions;

10 programmed means for removing a first gene from a cluster when
11 the first gene is also in another cluster which has another gene
12 with a higher match fraction with the first gene than any of the
13 genes in the cluster have with the first gene;

14 programmed means for repeating the removing step until all
15 genes are listed in only one cluster.

1 11. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles using output from a similar sequences algorithm
4 that is a time and intensity invariant correlation function
5 comprising:

6 programmed means for providing a gene expression profile data
7 set as input to programmed means embodying a similar sequences
8 algorithm that is a time and intensity invariant correlation
9 function to obtain a data set of gene expression pairs and a match
10 fraction for each pair as output from the programmed means embodying
11 a similar sequences algorithm;

12 programmed means for listing the gene expression pairs in
13 clusters by their match fractions;

14 programmed means for removing a first gene from a cluster when
15 the first gene is also in another cluster which has another gene
16 with a higher match fraction with the first gene than any of the
17 genes in the cluster have with the first gene;

18 programmed means for repeating the removing step until all
19 genes are listed in only one cluster.

1 12. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;

8 programmed means for listing gene expression pairs in clusters
9 by their match fractions;

10 programmed means for removing a first gene from a first cluster
11 when the first gene is also in a second cluster which has another
12 gene with a higher match fraction with the first gene than any of
13 the genes in the first cluster have with the first gene, but;

14 programmed means for retaining the first gene in the first
15 cluster and removing the first gene from the second cluster when the
16 difference between the highest match fraction of the first gene with
17 a gene in the first cluster and the highest match fraction of the
18 first gene with a gene in the second cluster is less than a minimum
19 difference threshold and the number of subsequences represented in
20 the similar gene pair having the highest match fraction in the first
21 cluster is higher than the number of subsequences represented in the
22 similar gene pair having the highest match fraction in the second
23 cluster;

24 programmed means for repeating the removing step until all
25 genes are listed in only one cluster.

1 13. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;

8 programmed means for choosing a threshold match fraction;

9 programmed means for listing gene expression pairs in clusters
10 by their match fractions above the threshold;

11 programmed means for adding each gene not already in a cluster
12 to a cluster having another gene having a highest match fraction
13 with the each gene without regard of the threshold;

14 programmed means for removing a first gene from a cluster when
15 the first gene is also in another cluster which has another gene
16 with a higher match fraction with the first gene than any of the
17 genes in the cluster have with the first gene;

18 programmed means for repeating the removing step until all
19 genes are listed in only one cluster.

1 14. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:

4 programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;

8 programmed means for choosing a threshold match fraction;

9 programmed means for listing gene expression pairs in clusters
10 by their match fractions above the threshold;

11 programmed means for adding each gene not already in a cluster
12 to a cluster having another gene having a highest match fraction
13 disregarding the threshold with the each gene;

14 programmed means for removing a first gene from a first cluster
15 when the first gene is also in a second cluster which has another
16 gene with a higher match fraction with the first gene than any of
17 the genes in the first cluster have with the first gene, but;

18 programmed means for retaining the first gene in the first
19 cluster and removing the first gene from the second cluster when the
20 difference between the highest match fraction of the first gene with
21 a gene in the first cluster and the highest match fraction of the
22 first gene with a gene in the second cluster is less than a minimum
23 difference threshold and the number of subsequences represented in
24 the similar gene pair having the highest match fraction in the first
25 cluster is higher than the number of subsequences represented in the
26 similar gene pair having the highest match fraction in the second
27 cluster;

28 programmed means for repeating the removing and retaining steps
29 until all genes are listed in only one cluster.

1 15. A program product having computer readable code stored on a
2 recordable media for determining similarity between genes comprising
3 the steps of:

4 programmed means for listing genes to be compared by their gene
5 expression profiles;

6 programmed means for processing the listed gene expression
7 profiles with a similar sequences algorithm that is a time and
8 intensity invariant correlation function to obtain a data set of
9 gene expression pairs and a match fraction for each pair;

10 programmed means for choosing a threshold match fraction;

11 programmed means for creating a null set $G(0)$ to hold genes
12 accounted for;

13 programmed means for assigning genes i and j to cluster 1 if
14 they have a match fraction greater than the threshold;

15 programmed means for assigning gene k to cluster 1 if it has a
16 match fraction greater than the threshold with either gene i or gene
17 j ;

18 programmed means for assigning genes k and l to cluster 2 if
19 they have a match fraction greater than the threshold and if both
20 gene k and gene l do not have match fractions above the threshold
21 with either gene i or gene j ;

22 programmed means for removing a first gene from a cluster when
23 another cluster has another gene with a higher match fraction with
24 the first gene;

25 programmed means for repeating the removing step until all
26 genes are listed in only one cluster.

1 16. A program product having computer readable code stored on a
2 recordable media for determining similarity between genes comprising
3 the steps of:

4 programmed means for listing genes to be compared by their gene
5 expression profiles;

6 programmed means for processing the listed gene expression
7 profiles with a similar sequences algorithm that is a time and
8 intensity invariant correlation function to obtain a data set of
9 gene expression pairs and a match fraction for each pair;

10 programmed means for choosing a threshold match fraction;

11 programmed means for creating a null set $G(0)$ to hold genes
12 accounted for;

13 programmed means for assigning genes i and j to cluster 1 if
14 they have a match fraction greater than the threshold;

15 programmed means for assigning gene k to cluster 1 if it has a
16 match fraction greater than the threshold with either gene i or gene
17 j ;

18 programmed means for assigning genes k and l to cluster 2 if
19 they have a match fraction greater than the threshold and if both
20 gene k and gene l do not have match fractions above the threshold
21 with either gene i or gene j ;

22 programmed means for removing a first gene from a cluster when
23 another cluster has another gene with a higher match fraction with
24 the first gene, unless the another gene requires a larger number of
25 subsequences to achieve similarity with the first gene;

26 programmed means for repeating the removing step until all
27 genes are listed in only one cluster.
28

1 20. In a method of determining similarity between portions of
2 gene expression profiles which includes processing a number of gene
3 expression profiles using a computer with a similar sequences
4 algorithm that is a time and intensity invariant correlation
5 function to obtain a data set of gene expression pairs and a match
6 fraction for each pair, the improvement comprising the steps of:
7 listing gene expression pairs in clusters by their match
8 fractions;
9 removing a first gene from a cluster when another cluster has
10 another gene with a higher match fraction with the first gene,
11 unless the another gene requires a larger number of subsequences to
12 achieve similarity with the first gene;
13 repeating the removing step until all genes are listed in only
14 one cluster;
15 providing output of the listing of clusters of gene
16 expression profiles.

17

1 **(IX) Evidence Appendix**

2 No evidence is being submitted in this appeal.

1 **(X) Related Proceedings Appendix**

2 None.